CLOSER – UK LLC webinar: a step-by-step guide to applying and accessing linked data in the UK LLC Trusted Research Environment

Introduction

Robin Flaig
Deputy Director
5 April 2022

AGENDA

Introduction	Robin Flaig
Application process	Stela McLachlan
Data Request	Sammy Berman

Break for questions

Data Access	Rich Thomas
Public Involvement	Kirsteen Campbell
Legal Documentation	Jacqui Oakley

Break for questions

Conclusions and summary Robin Flaig

A collaboration to benefit public health

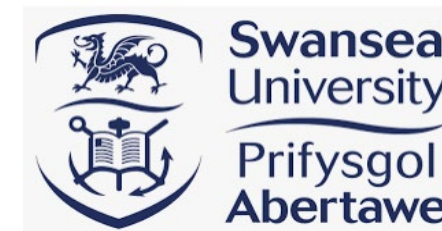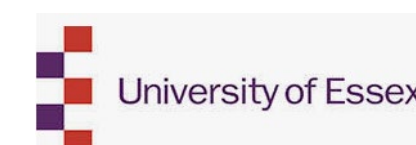UK LLC led by the Universities of Edinburgh and Bristol

In collaboration with

A work package in the COVID-19 Longitudinal Health & Wellbeing National Core Study

UK Longitudinal Linkage Collaboration

A four-nations infrastructure for secure cross-cohort analysis

Align complex and divergent governance frameworks into a streamlined and predictable linkage/access route

Develop a sustainable four-nations data science infrastructure

A new way of working for record linkage

Watch our animation: https://youtu.be/zFkvpKD3jvs

UK LLC Design Decisions:

Interdisciplinary

Four nations approach

Pre-pandemic & pandemic data

Trusted Research Environment

Five Safes access model

De-personalised data only

Distributed governance

Transparent in its operations

Public Involvement programme

A globally unique cross-cutting resource

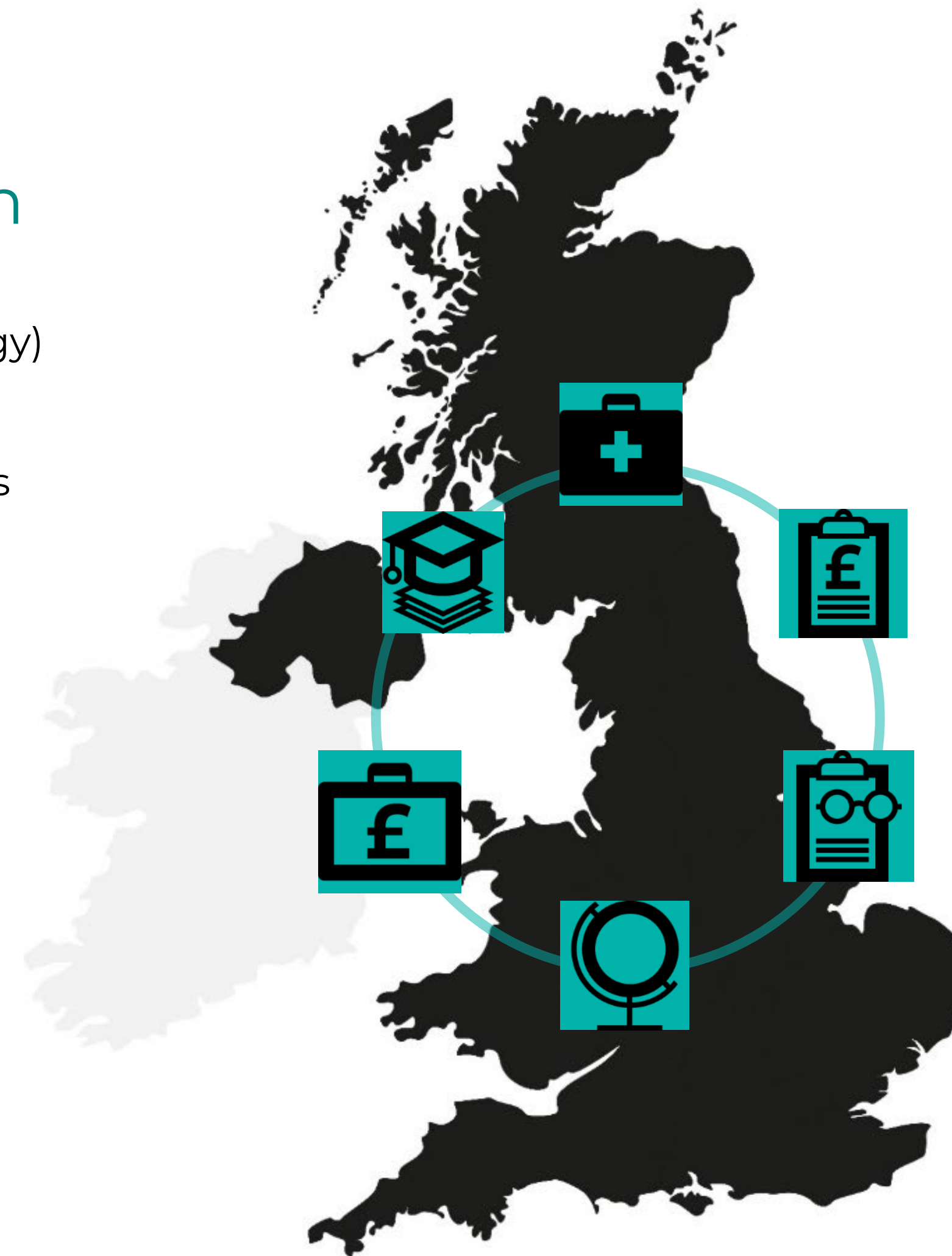## 20+ Longitudinal Population Studies
- COVID-19 collections (Questionnaires, Serology)
- SES & demographic
- Baseline physical & mental health
- Baseline family, SES and life-course indicators
- Genetics

NHS COVID-19 datasets
- GP Data
- Vaccinations
- COVID-19 test data (Pillar 1-3)
- Accident & Emergency
- CHESS

Wider NHS datasets
- Hospital Inpatient data
- Cancer & mortality registers
- Community Mental Health (MHSDS, IAPT)
- Prescribing data

Administrative records
- HMRC
- DWP
- DfE

Negotiating data access

Environmental & neighbourhood
- Air pollution
- Noise
- Greenspace
- SES & demographic
- Service provision

Being modelled

Public Involvement

4 touchpoints for involving study participants and members of the public:

Strategic Advisory Committee

Data Access Public Review Panel

Public Advisory Group

Public Involvement Network

- Strategy and future collaborations
- Involve in data access process
- Assess remit and scope of public involvement
- Public involvement policy development
- Mutual ways of working
- Access wider public involvement networks
- Acknowledge and communicate public involvement
- Measure impact
- Training and guidance

The Minimum Viable Product

LLE UK LLC

COVID-19 LONGITUDINAL HEALTH & WELLBEING
NATIONAL CORE STUDY

> 20+ longitudinal studies
> Novel NHS Digital data pipeline
> Public contributors & application panel
> Nine applications through beta-testing phase

Year 1

Accessible data from:

ALSPAC: Avon Longitudinal Study of Parents and Children

1970 British Cohort Study

Born in Bradford

ELSA: English Longitudinal Study of Ageing

EPIC-Norfolk: The European Prospective Investigation into Cancer (EPIC) Norfolk Study

EXCEED: Extended Cohort for E-health, Environment & DNA

The Fenland Study

Generation Scotland

GLAD: Genetic Links to Anxiety and Depression Study

The Millennium Cohort Study

NCDS58: 1958 National Child Development Study

Next Steps

NICOLA: Northern Ireland Cohort for the Longitudinal Study of Ageing (no linked data)

NIHR BioResource_COPING: National Institute of Health Research BioResource, COVID-19 Psychiatry and Neurological Genetics Study

NSHD46: MRC National Survey of Health and Development Cohort/1946 Birth Cohort

SABRE (no linked data)

TRACK-COVID Study

TwinsUK

Understanding Society - the UK Household Longitudinal Study

Keywords most commonly used to describe the LPS data

Linked Data – currently available

NHS Digital

- Primary Care GPES: Data for Pandemic Planning and Research
- COVID-19 Vaccination Status & COVID-19 Adverse Reaction
- NPex (pillar 2) Mental Health Services Data Set
- Medicines Dispensed in Primary Care (NHSBSA data)
- Hospital Episode Statistics: Admitted Patient Care, Critical Care, Outpatients, A&E
- Demographics data set
- COVID-19 SARI-Watch (formerly CHESS) COVID-19
- Second Generation Surveillance System (Pillars 1 & 2)
- COVID-19 UK Non-hospital Antibody Testing Results (Pillar 3) Civil Registration – Deaths (England & Wales)
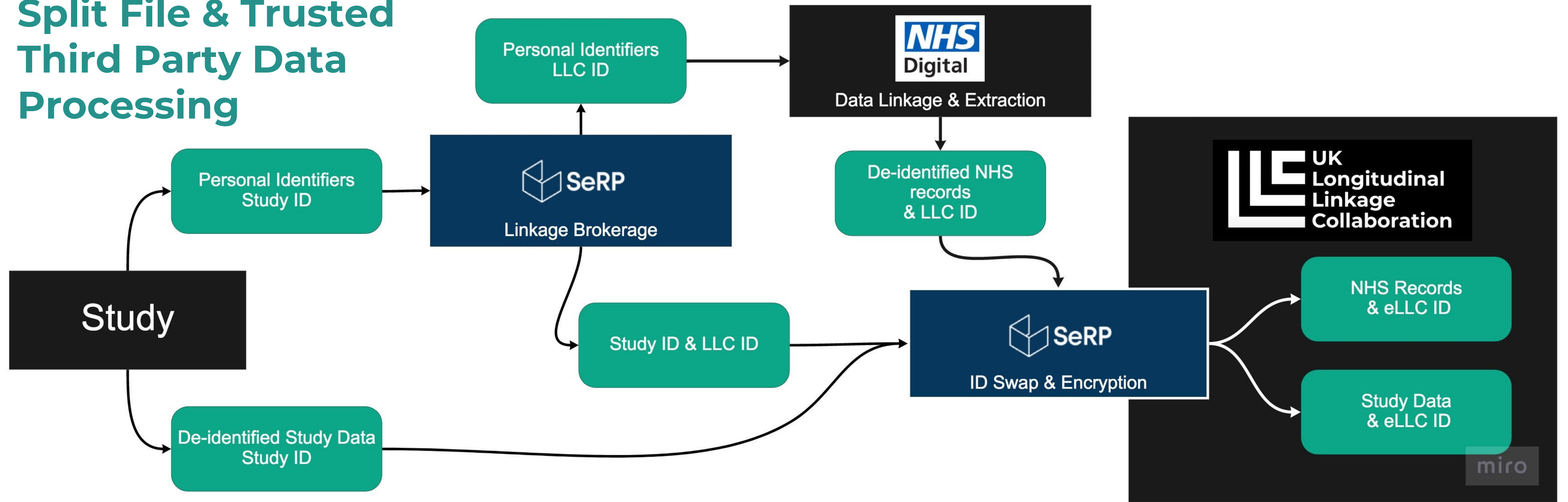- Cancer Registration Data (England & Wales)

Generation Scotland

- PHS - NHS Vaccination data
- Electronic Communication of Surveillance in Scotland - testing
- PHS - NHS Prescription data
- PHS - Acute Inpatient and Day Case
- Civil Registration - Deaths

UK LLC Trusted Research Environment (TRE)

UK LLC

COVID-19 LONGITUDINAL HEALTH & WELLBEING
NATIONAL CORE STUDY

Split File & Trusted Third Party Data Processing

Study

Personal Identifiers
Study ID

De-identified Study Data
Study ID

SeRP
Linkage Brokerage

Personal Identifiers
LLC ID

Study ID & LLC ID

NHS Digital
Data Linkage & Extraction

De-identified NHS records
& LLC ID

SeRP
ID Swap & Encryption

UK Longitudinal Linkage Collaboration

NHS Records
& eLLC ID

Study Data
& eLLC ID

miro

Functionally Anonymous TRE

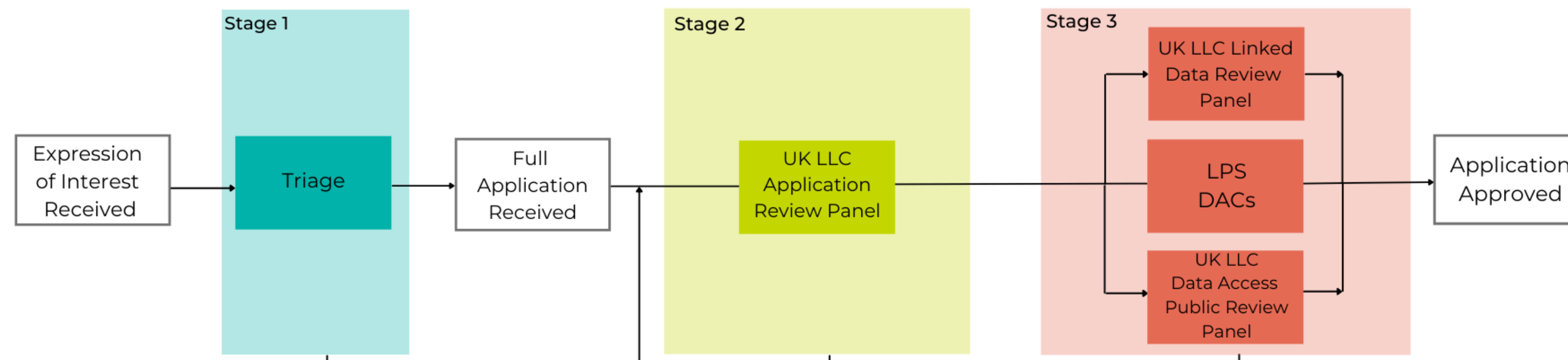# UK LLC: efficient and predictable data access for analysts

Data linkage made easy and predictable

UK LLC 'Five Safes' Application Framework

- UK LLC distributes applications to contributing studies
- UK LLC has delegated review authority from NHS Digital
- COVID-19 research only, UK based analysts
- Public benefit with public review
- Requirement for reproduceable research

Apply for access via portal at the HDRUK Gateway (https://www.healthdatagateway.org/)
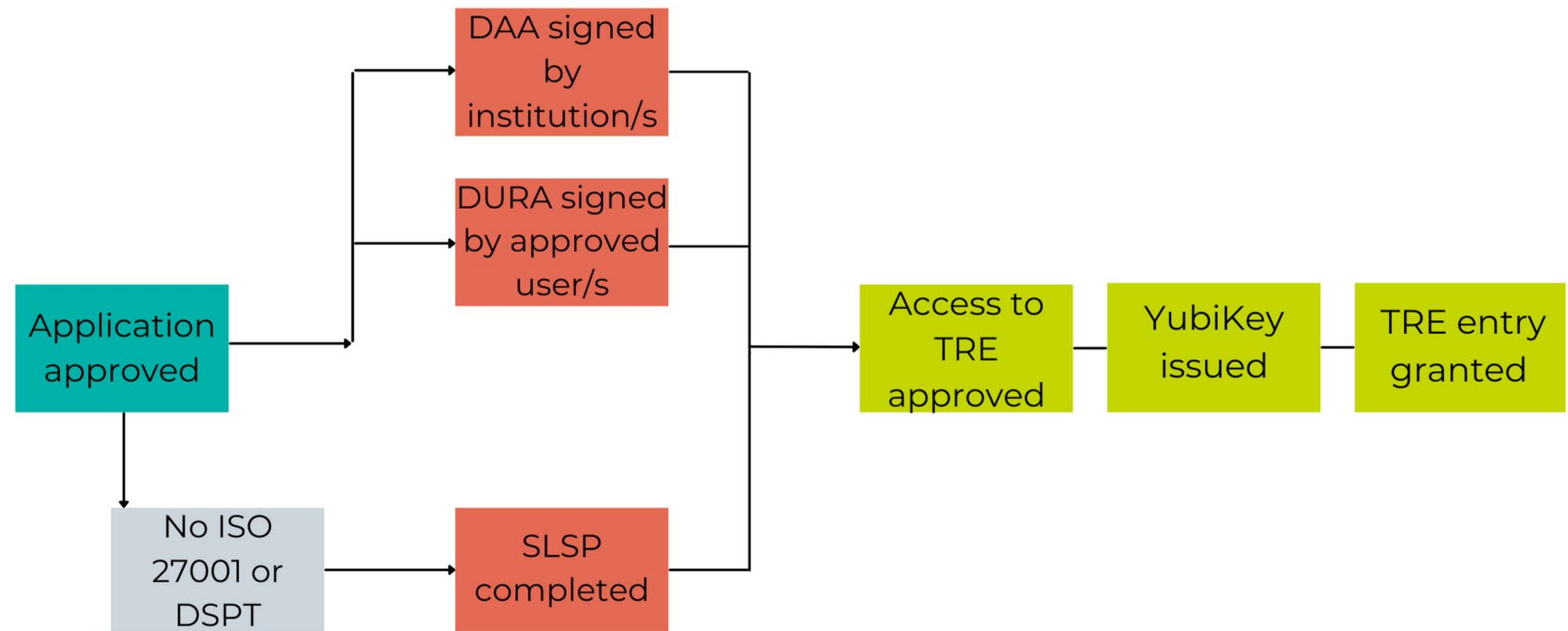
- All researchers need to be ONS Approved Researchers & trained
- Data Access Agreement
- Light touch research user IT security framework as heavy lifting done at the UKLLC (ISO27001)

UK LLC Access – after approval

- Each research institution must sign the Data Access Agreement (DAA) and have IG covered

- Each researcher must sign a Data User Responsibilities Agreement (DURA)

- Once one institution and researcher have signed the DAA and DURA data can be provisioned
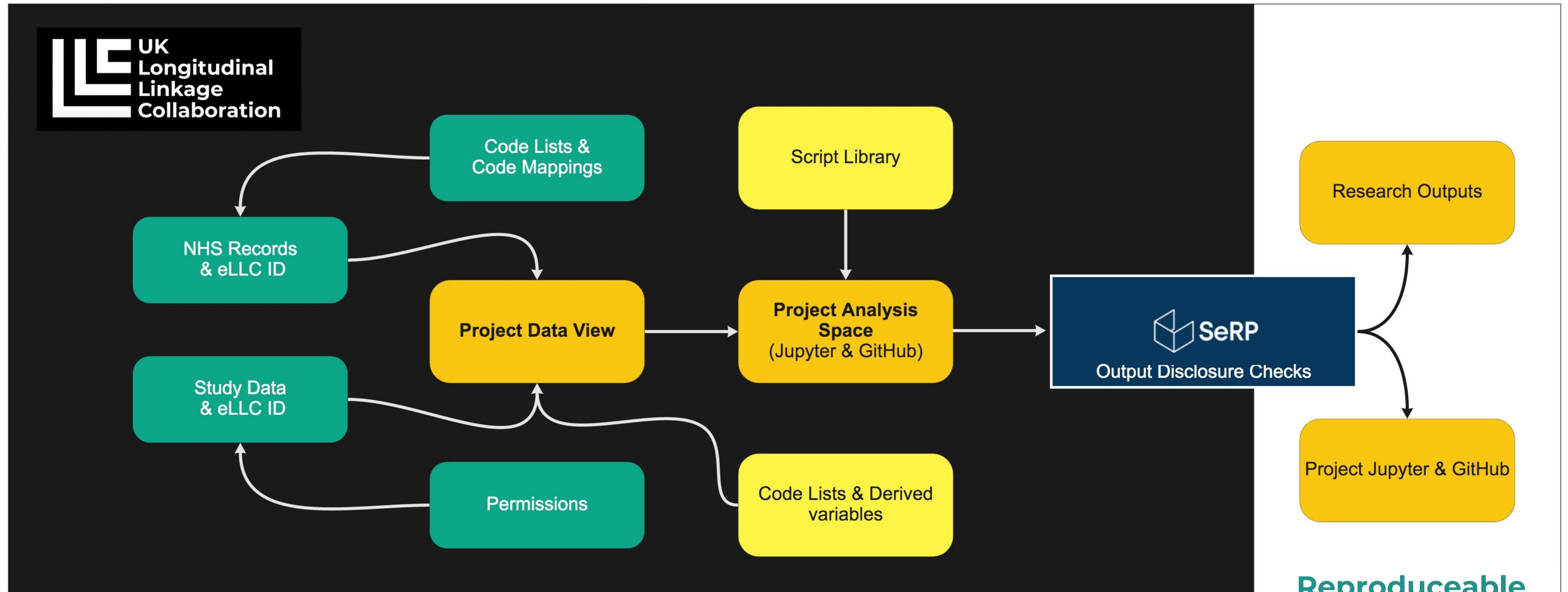
- Amendments to approved applications can be made

User experience & reproduceable research

All processing *within* the TRE

Safe Outputs

Reproduceable Research

Future plans

User experience & reproducible research

Enhance with administrative records

Efficient and predictable data access for analysts

Expanding beyond Covid-19 research to become a general research resource.